

Optimization and Simulation

Drawing from distributions

Michel Bierlaire

Transport and Mobility Laboratory
School of Architecture, Civil and Environmental Engineering
Ecole Polytechnique Fédérale de Lausanne



Outline

- 1 Discrete distributions
- 2 Continuous distributions
- 3 Transforming draws
- 4 Monte-Carlo integration
- 5 Summary
- 6 Appendix

Discrete distributions

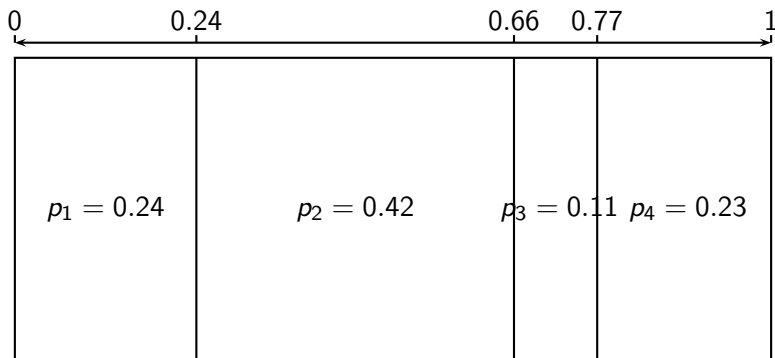
- Let X be a discrete r.v. with pmf:

$$P(X = x_i) = p_i, \quad i = 0, \dots,$$

where $\sum_i p_i = 1$.

- The support can be finite or infinite.
- We know how to draw from $U(0, 1)$.
- How can we draw from X ?

Inverse Transform Method: illustration



Discrete distributions

Inverse transform method

- 1 Let r be a draw from $U(0, 1)$.
- 2 Initialize $k = 0$, $p = 0$.
- 3 $p = p + p_k$.
- 4 If $r < p$, set $X = x_k$ and stop.
- 5 Otherwise, set $k = k + 1$ and go to step 3.

Discrete distributions

Acceptance-rejection

- Attributed to von Neumann.
- Mostly useful with continuous distributions.
- We want to draw from X with pmf p_i .
- We know how to draw from Y with pmf q_i .

Define a constant $c \geq 1$ such that

$$\frac{p_i}{q_i} \leq c \quad \forall i \text{ s.t. } p_i > 0.$$

Algorithm

- 1 Draw y from Y
- 2 Draw r from $U(0, 1)$
- 3 If $r < \frac{p_y}{cq_y}$, return $x = y$ and stop. Otherwise, start again.

Acceptance-rejection: analysis

Probability to be accepted during a given iteration

$$\begin{aligned}
 P(Y = y, \text{accepted}) &= P(Y = y) P(\text{accepted} | Y = y) \\
 &= q_y \quad p_y / cq_y \\
 &= \frac{p_y}{c}
 \end{aligned}$$

Probability to be accepted

$$\begin{aligned}
 P(\text{accepted}) &= \sum_y P(\text{accepted} | Y = y) P(Y = y) \\
 &= \sum_y \frac{p_y}{cq_y} q_y \\
 &= 1/c.
 \end{aligned}$$

Probability to draw x at iteration n

$$P(X = x | n) = \left(1 - \frac{1}{c}\right)^{n-1} \frac{p_x}{c}$$

Acceptance-rejection: analysis

$$\begin{aligned}
 P(X = x) &= \sum_{n=1}^{+\infty} P(X = x|n) \\
 &= \sum_{n=1}^{+\infty} \left(1 - \frac{1}{c}\right)^{n-1} \frac{p_x}{c} \\
 &= c \frac{p_x}{c} \\
 &= p_x.
 \end{aligned}$$

Reminder: geometric series

$$\sum_{n=0}^{+\infty} x^n = \frac{1}{1-x}$$

Acceptance-rejection: analysis

Remarks

- Average number of iterations: c
- The closer c is to 1, the closer the pmf of Y is to the pmf of X .

Outline

- 1 Discrete distributions
- 2 Continuous distributions**
- 3 Transforming draws
- 4 Monte-Carlo integration
- 5 Summary
- 6 Appendix

Continuous distributions

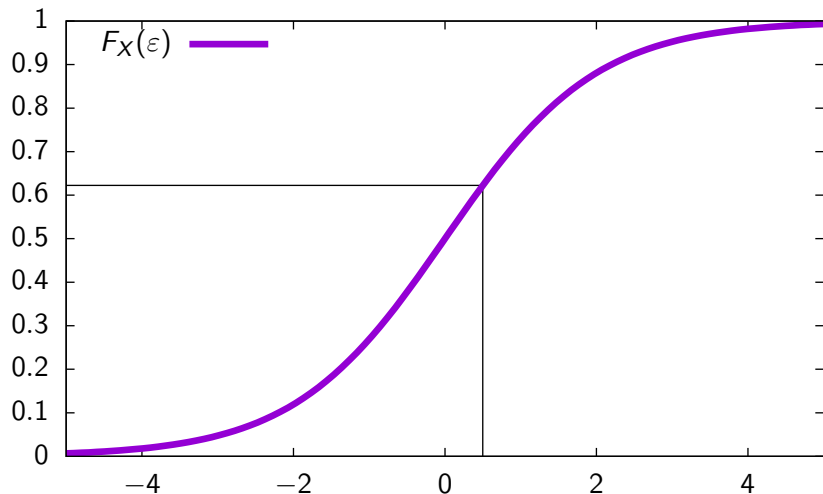
Inverse Transform Method

- Let X be a continuous r.v. with CDF $F_X(\varepsilon)$
- Draw r from a uniform $U(0, 1)$
- Generate $F_X^{-1}(r)$.

Motivation

- F_X is monotonically increasing
- It implies that $\varepsilon_1 \leq \varepsilon_2$ is equivalent to $F_X(\varepsilon_1) \leq F_X(\varepsilon_2)$.

Inverse Transform Method



Inverse Transform Method

More formally

- Denote $F_U(\varepsilon) = \varepsilon$ the CDF of the r.v. $U(0, 1)$
- Let G be the distribution of the r.v. $F_X^{-1}(U)$

$$\begin{aligned}G(\varepsilon) &= \Pr(F_X^{-1}(U) \leq \varepsilon) \\&= \Pr(F_X(F_X^{-1}(U)) \leq F_X(\varepsilon)) \\&= \Pr(U \leq F_X(\varepsilon)) \\&= F_U(F_X(\varepsilon)) \\&= F_X(\varepsilon)\end{aligned}$$

Inverse Transform Method

Examples: let r be a draw from $U(0, 1)$

Name	$F_X(\varepsilon)$	Draw
Exponential(b)	$1 - e^{-\varepsilon/b}$	$-b \ln r$
Logistic(μ, σ)	$1 / (1 + \exp(-(\varepsilon - \mu) / \sigma))$	$\mu - \sigma \ln(\frac{1}{r} - 1)$
Power(n, σ)	$(\varepsilon / \sigma)^n$	$\sigma r^{1/n}$

Note

The CDF is not always available (e.g. normal distribution).

Continuous distributions

Rejection Method

- We want to draw from X with pdf f_X .
- We know how to draw from Y with pdf f_Y .

Define a constant $c \geq 1$ such that

$$\frac{f_X(\varepsilon)}{f_Y(\varepsilon)} \leq c \quad \forall \varepsilon$$

Algorithm

- 1 Draw y from Y
- 2 Draw r from $U(0, 1)$
- 3 If $r < \frac{f_X(y)}{cf_Y(y)}$, return $x = y$ and stop. Otherwise, start again.

Rejection Method: example

Draw from a normal distribution

- Let $\bar{X} \sim N(0, 1)$ and $X = |\bar{X}|$
- Probability density function: $f_X(\varepsilon) = \frac{2}{\sqrt{2\pi}} e^{-\varepsilon^2/2}$, $0 < \varepsilon < +\infty$
- Consider an exponential r.v. with pdf $f_Y(\varepsilon) = e^{-\varepsilon}$, $0 < \varepsilon < +\infty$
- Then

$$\frac{f_X(\varepsilon)}{f_Y(\varepsilon)} = \frac{2}{\sqrt{2\pi}} e^{\varepsilon - \varepsilon^2/2}$$

- The ratio takes its maximum at $\varepsilon = 1$, therefore

$$\frac{f_X(\varepsilon)}{f_Y(\varepsilon)} \leq \frac{f_X(1)}{f_Y(1)} = \sqrt{2e/\pi} \approx 1.315.$$

- Rejection method, with $\frac{f_X(\varepsilon)}{cf_Y(\varepsilon)} = \frac{1}{\sqrt{e}} e^{\varepsilon - \varepsilon^2/2} = e^{\varepsilon - \frac{\varepsilon^2}{2} - \frac{1}{2}} = e^{-\frac{(\varepsilon-1)^2}{2}}$

Rejection Method: example

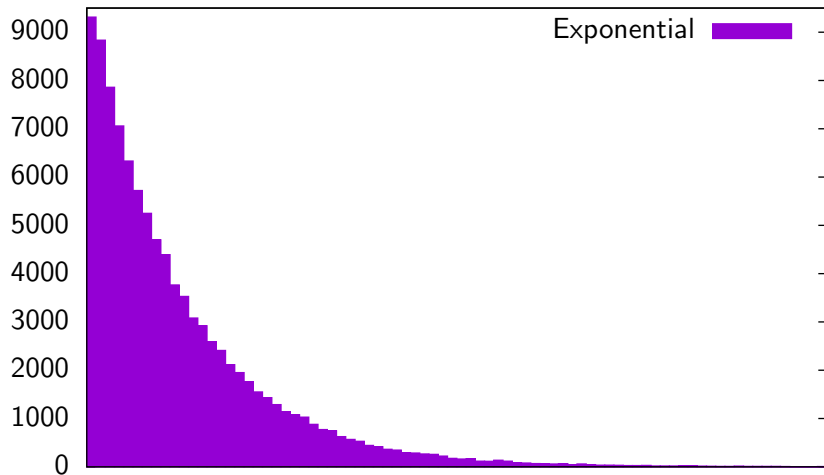
Algorithm: draw from a normal

- 1 Draw r from $U(0, 1)$
- 2 Let $y = -\ln(1 - r)$ (draw from the exponential)
- 3 Draw s from $U(0, 1)$
- 4 If $s < e^{-\frac{(y-1)^2}{2}}$ return $x = y$ and go to step 5. Otherwise, go to step 1.
- 5 Draw t from $U(0, 1)$.
- 6 If $t \leq 0.5$, return x . Otherwise, return $-x$.

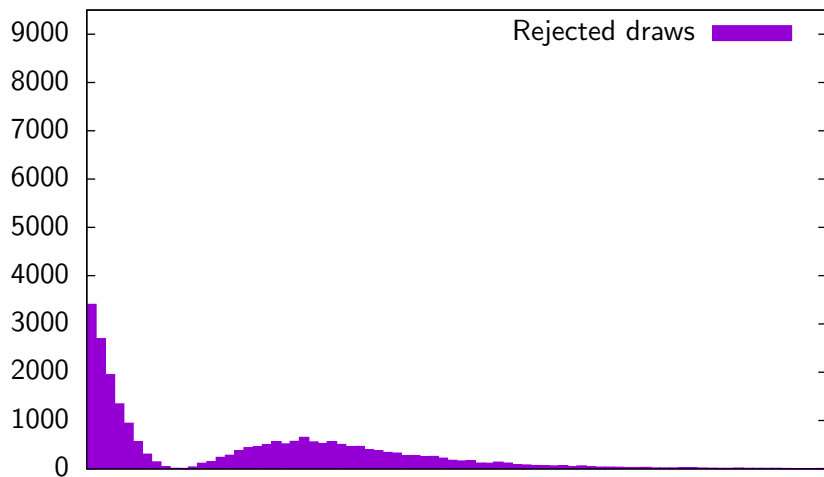
Note

This procedure can be improved. See [Ross, 2006, Chapter 5].

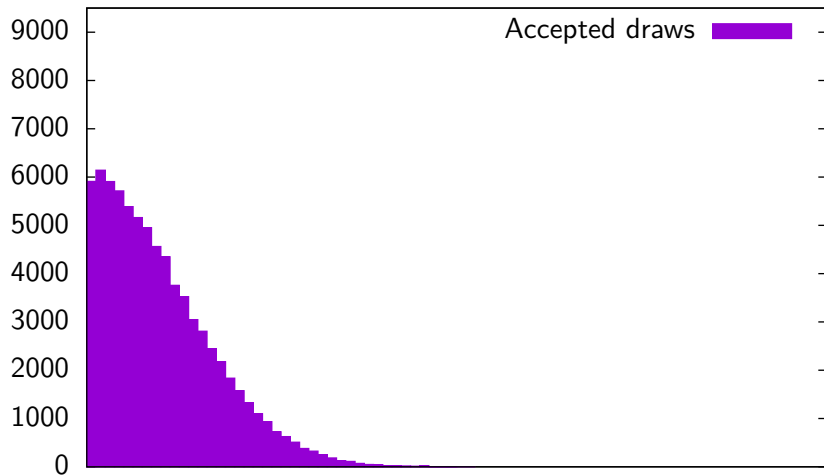
Draws from the exponential



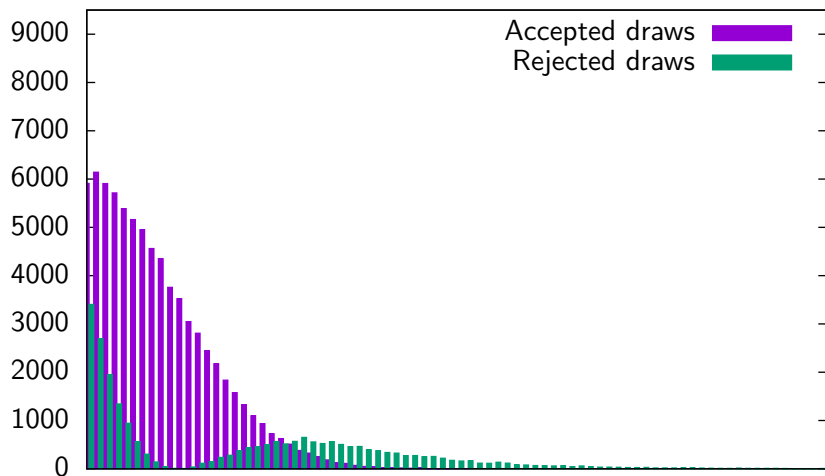
Rejected draws



Accepted draws



Rejected and accepted draws



Drawing from the standard normal distribution

- Accept/reject algorithm is not efficient
- Polar method: no rejection (see appendix)

Transformations of standard normal

- If r is a draw from $N(0, 1)$, then

$$s = br + a$$

is a draw from $N(a, b^2)$

- If r is a draw from $N(a, b^2)$, then

$$e^r$$

is a draw from a log normal $LN(a, b^2)$ with mean

$$e^{a+(b^2/2)}$$

and variance

$$e^{2a+b^2}(e^{b^2} - 1)$$

Multivariate normal

- If r_1, \dots, r_n are independent draws from $N(0, 1)$, and

$$r = \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix}$$

- then

$$s = a + Lr$$

is a vector of draws from the n -variate normal $N(a, LL^T)$, where

- L is lower triangular, and
- LL^T is the Cholesky factorization of the variance-covariance matrix

Multivariate normal

Example:

$$L = \begin{pmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{pmatrix}$$

$$s_1 = l_{11}r_1$$

$$s_2 = l_{21}r_1 + l_{22}r_2$$

$$s_3 = l_{31}r_1 + l_{32}r_2 + l_{33}r_3$$

Outline

- 1 Discrete distributions
- 2 Continuous distributions
- 3 Transforming draws**
- 4 Monte-Carlo integration
- 5 Summary
- 6 Appendix

Transforming draws

Method

- Consider draws from the following distributions:
 - normal: $N(0, 1)$ (draws denoted by ξ below)
 - uniform: $U(0, 1)$ (draws denoted by r below)
- Draws R from other distributions are obtained from nonlinear transforms.

Lognormal(a,b)

$$f(x) = \frac{1}{xb\sqrt{2\pi}} \exp\left(\frac{-(\ln x - a)^2}{2b^2}\right) \quad R = e^{a+b\xi}$$

Transforming draws

Cauchy(a,b)

$$f(x) = \left(\pi b \left(1 + \left(\frac{x-a}{b} \right)^2 \right) \right)^{-1} \quad R = a + b \tan \left(\pi \left(r - \frac{1}{2} \right) \right)$$

$\chi^2(a)$ (a integer)

$$f(x) = \frac{x^{(a-2)/2} e^{-x/2}}{2^{a/2} \Gamma(a/2)} \quad R = \sum_{j=1}^a \xi_j^2$$

Erlang(a,b) (b integer)

$$f(x) = \frac{(x/a)^{b-1} e^{-x/a}}{a(b-1)!} \quad R = -a \sum_{j=1}^b \ln r_j$$

Transforming draws

Exponential(a)

$$F(x) = 1 - e^{-x/a} \quad R = -a \ln r$$

Extreme Value(a,b)

$$F(x) = 1 - \exp\left(-e^{-(x-a)/b}\right) \quad R = a - b \ln(-\ln r)$$

Logistic(a,b)

$$F(x) = \left(1 + e^{-(x-a)/b}\right)^{-1} \quad R = a + b \ln\left(\frac{r}{1-r}\right)$$

Transforming draws

Pareto(a,b)

$$F(x) = 1 - \left(\frac{a}{x}\right)^b \quad R = a(1 - r)^{-1/b}$$

Standard symmetrical triangular distribution

$$f(x) = \begin{cases} 4x & \text{if } 0 \leq x \leq 1/2 \\ 4(1-x) & \text{if } 1/2 \leq x \leq 1 \end{cases} \quad R = \frac{r_1 + r_2}{2}$$

Weibull(a,b)

$$F(x) = 1 - e^{-\left(\frac{x}{a}\right)^b} \quad R = a(-\ln r)^{1/b}$$

Outline

- 1 Discrete distributions
- 2 Continuous distributions
- 3 Transforming draws
- 4 Monte-Carlo integration**
- 5 Summary
- 6 Appendix

Monte-Carlo integration

Expectation

- X r.v. on $[a, b]$, $a \in \mathbb{R} \cup \{-\infty\}$, $b \in \mathbb{R} \cup \{+\infty\}$
- Expectation of X :

$$E[X] = \int_a^b x f_X(x) dx.$$

- If $g : \mathbb{R} \rightarrow \mathbb{R}$ is a function, then

$$E[g(X)] = \int_a^b g(x) f_X(x) dx.$$

Monte-Carlo integration

Simulation

$$E[g(X)] \approx \frac{1}{R} \sum_{r=1}^R g(x_r).$$

Approximating the integral

$$\int_a^b g(x) f_X(x) dx = \lim_{R \rightarrow \infty} \frac{1}{R} \sum_{r=1}^R g(x_r).$$

so that

$$\int_a^b g(x) f_X(x) dx \approx \frac{1}{R} \sum_{r=1}^R g(x_r).$$

Monte-Carlo integration

Calculating $I = \int_a^b g(x)dx$

- Select X with known pdf f_X .
- Generate R draws x_r , $r = 1, \dots, R$ from X ;
- Calculate

$$I \approx \hat{I} = \frac{1}{R} \sum_{r=1}^R \frac{g(x_r)}{f_X(x_r)}.$$

Monte-Carlo integration

Approximation error

- Sample variance:

$$V_R = \frac{1}{R-1} \sum_{r=1}^R \left(\frac{g(x_r)}{f_X(x_r)} - \hat{I} \right)^2.$$

- By simulation: as

$$\text{Var}[g(X)] = E[g(X)^2] - E[g(x)]^2,$$

we have

$$V_R \approx \frac{1}{R} \sum_{r=1}^R \frac{g(x_r)^2}{f_X(x_r)} - \hat{I}^2.$$

Monte-Carlo integration

Approximation error

95% confidence interval: $[\hat{I} - 1.96e_R \leq I \leq \hat{I} + 1.96e_R]$ where

$$e_R = \sqrt{\frac{V_R}{R}}.$$

Monte-Carlo integration

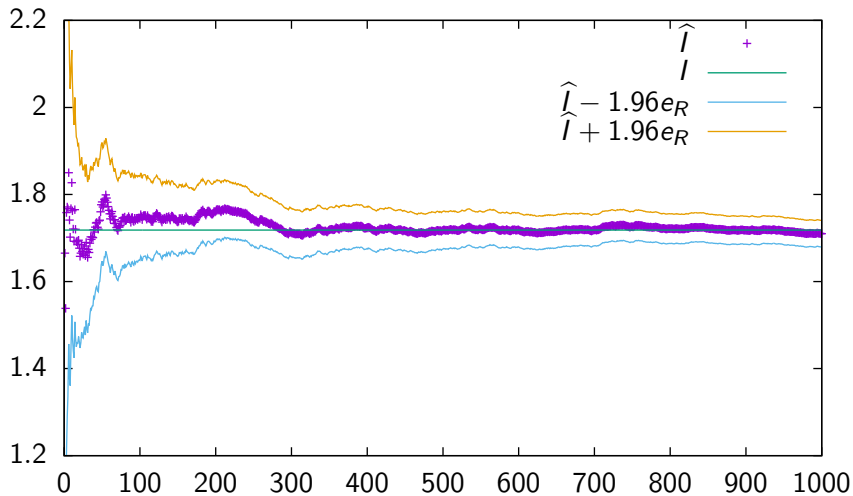
Example

$$\int_0^1 e^x dx = e - 1 = 1.7183$$

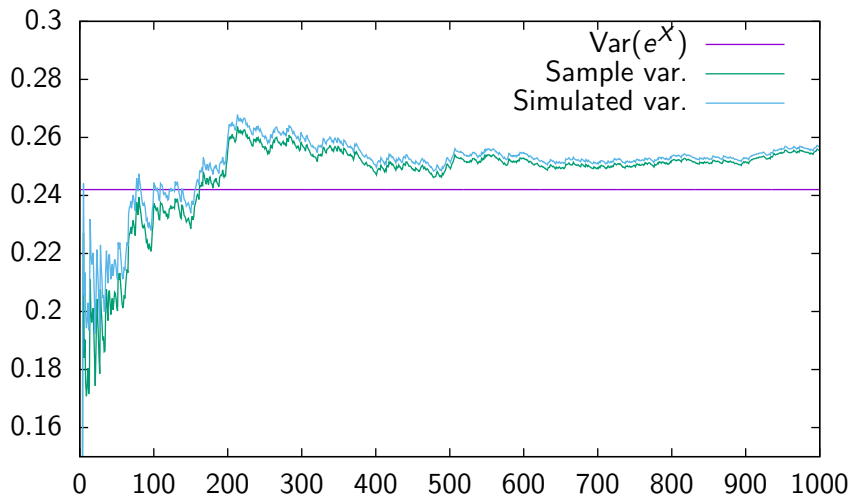
- Random variable X uniformly distributed ($f_X(\varepsilon) = 1$)
- $g(X) = e^X$
- $\text{Var}(e^X) = \frac{e^2-1}{2} - (e-1)^2 = 0.2420$

R	10	100	1000
\hat{I}	1.8270	1.7707	1.7287
Sample variance	0.1607	0.2125	0.2385
Simulated variance	0.1742	0.2197	0.2398

Monte-Carlo integration



Monte-Carlo integration



Outline

- 1 Discrete distributions
- 2 Continuous distributions
- 3 Transforming draws
- 4 Monte-Carlo integration
- 5 Summary**
- 6 Appendix

Summary

- Draws from uniform distribution: available in any programming language
- Inverse transform method: requires the pmf or the CDF.
- Accept-reject: needs a “similar” r.v. easy to draw from.
- Transforming uniform and normal draws.
- First application: Monte-Carlo integration.

Outline

- 1 Discrete distributions
- 2 Continuous distributions
- 3 Transforming draws
- 4 Monte-Carlo integration
- 5 Summary
- 6 Appendix**

Uniform distribution: $X \sim U(a, b)$

pdf

$$f_X(x) = \begin{cases} 1/(b-a) & \text{if } a \leq x \leq b, \\ 0 & \text{otherwise.} \end{cases}$$

CDF

$$F_X(x) = \begin{cases} 0 & \text{if } x \leq a, \\ (x-a)/(b-a) & \text{if } a \leq x \leq b, \\ 1 & \text{if } x \geq b. \end{cases}$$

Mean, median

$$(a+b)/2$$

Variance

$$(b-a)^2/12$$

Normal distribution: $X \sim N(a, b)$

pdf

$$f_X(x) = \frac{1}{b\sqrt{2\pi}} \exp\left(-\frac{(x-a)^2}{2b^2}\right)$$

CDF

$$F_X(x) = \int_{-\infty}^x f_X(t) dt.$$

Mean, median

a

Variance

b^2

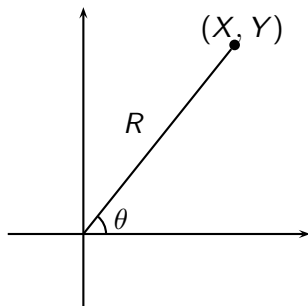
The polar method

Draw from a normal distribution

- Let $X \sim N(0, 1)$ and $Y \sim N(0, 1)$ independent
- pdf:

$$f(x, y) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} = \frac{1}{2\pi} e^{-(x^2+y^2)/2}.$$

- Let R and θ such that $R^2 = X^2 + Y^2$, and $\tan \theta = Y/X$.



The polar method

Change of variables (reminder)

- Let A be a multivariate r.v. distributed with pdf $f_A(a)$.
- Consider the change of variables $b = H(a)$ where H is bijective and differentiable
- Then $B = H(A)$ is distributed with pdf

$$f_B(b) = f_A(H^{-1}(b)) \left| \det \left(\frac{dH^{-1}(b)}{db} \right) \right|.$$

Here: $A = (X, Y)$, $B = (R^2, \theta) = (T, \theta)$

$$H^{-1}(b) = \begin{pmatrix} T^{\frac{1}{2}} \cos \theta \\ T^{\frac{1}{2}} \sin \theta \end{pmatrix} \quad \frac{dH^{-1}(b)}{db} = \begin{pmatrix} \frac{1}{2} T^{-\frac{1}{2}} \cos \theta & -T^{\frac{1}{2}} \sin \theta \\ \frac{1}{2} T^{-\frac{1}{2}} \sin \theta & T^{\frac{1}{2}} \cos \theta \end{pmatrix}$$

The polar method

$$H^{-1}(b) = \begin{pmatrix} T^{\frac{1}{2}} \cos \theta \\ T^{\frac{1}{2}} \sin \theta \end{pmatrix} \quad \frac{dH^{-1}(b)}{db} = \begin{pmatrix} \frac{1}{2} T^{-\frac{1}{2}} \cos \theta & -T^{\frac{1}{2}} \sin \theta \\ \frac{1}{2} T^{-\frac{1}{2}} \sin \theta & T^{\frac{1}{2}} \cos \theta \end{pmatrix}$$

Therefore,

$$\left| \det \left(\frac{dH^{-1}(b)}{db} \right) \right| = \frac{1}{2}.$$

and

$$f_B(T, \theta) = \frac{1}{2} \frac{1}{2\pi} e^{-T/2}, \quad 0 < T < +\infty, \quad 0 < \theta < 2\pi.$$

Product of

- an exponential with mean 2: $\frac{1}{2} e^{-T/2}$
- a uniform on $[0, 2\pi[$: $1/2\pi$

The polar method

Therefore

- R^2 and θ are independent
- R^2 is exponential with mean 2
- θ is uniform on $(0, 2\pi)$

Algorithm

- 1 Let r_1 and r_2 be draws from $U(0, 1)$.
- 2 Let $R^2 = -2 \ln r_1$ (draw from exponential of mean 2)
- 3 Let $\theta = 2\pi r_2$ (draw from $U(0, 2\pi)$)
- 4 Let

$$\begin{aligned} X &= R \cos \theta = \sqrt{-2 \ln r_1} \cos(2\pi r_2) \\ Y &= R \sin \theta = \sqrt{-2 \ln r_1} \sin(2\pi r_2) \end{aligned}$$

The polar method

Issue

Time consuming to compute sine and cosine

Solution

Generate directly the result of the sine and the cosine

- Draw a random point (s_1, s_2) in the circle of radius one centered at $(0, 0)$.
- **How?** Draw a random point in the square $[-1, 1] \times [-1, 1]$ and reject points outside the circle
- Let (R, θ) be the polar coordinates of this point.
- $R^2 \sim U(0, 1)$ and $\theta \sim U(0, 2\pi)$ are independent

$$\begin{aligned}R^2 &= s_1^2 + s_2^2 \\ \cos \theta &= s_1/R \\ \sin \theta &= s_2/R\end{aligned}$$

The polar method

Original transformation

$$\begin{aligned} X &= R \cos \theta = \sqrt{-2 \ln r_1} \cos(2\pi r_2) \\ Y &= R \sin \theta = \sqrt{-2 \ln r_1} \sin(2\pi r_2) \end{aligned}$$

Draw (s_1, s_2) in the circle

$$\begin{aligned} t &= s_1^2 + s_2^2 \\ X &= R \cos \theta = \sqrt{-2 \ln t} \frac{s_1}{\sqrt{t}} = s_1 \sqrt{\frac{-2 \ln t}{t}} \\ Y &= R \sin \theta = \sqrt{-2 \ln t} \frac{s_2}{\sqrt{t}} = s_2 \sqrt{\frac{-2 \ln t}{t}} \end{aligned}$$

The polar method

Algorithm

- 1 Let r_1 and r_2 be draws from $U(0, 1)$.
- 2 Define $s_1 = 2r_1 - 1$ and $s_2 = 2r_2 - 1$ (draws from $U(-1, 1)$).
- 3 Define $t = s_1^2 + s_2^2$.
- 4 If $t > 1$, reject the draws and go to step 1.
- 5 Return

$$x = s_1 \sqrt{\frac{-2 \ln t}{t}} \text{ and } y = s_2 \sqrt{\frac{-2 \ln t}{t}}.$$

Bibliography



Ross, S. M. (2006).
Simulation.
Elsevier, fourth edition.