What we learned untill now:

- Why do we need simulation
- What are pseudo random numbers
- How to simulate from discrete distribution
- How to simulate from (some) continuous distributions
    - Probability integral transform

Review: Sampling from discrete distributions

Let X be a stochastic variable with possible values $\{x_1, \dots, x_k\}$ and $\mathrm{P}(X = x_i) = \mathrm{p}_i$, $\sum_{i=1}^{k} p_i = 1$.

Define: $F_0 = 0, F_1 = p1, F_2 = p_1 + p_2, \dots, F_k = 1$

We can simulate value from F as:

$u \sim U[0, 1]$

for $i = 1, 2, \dots, k$ **do**

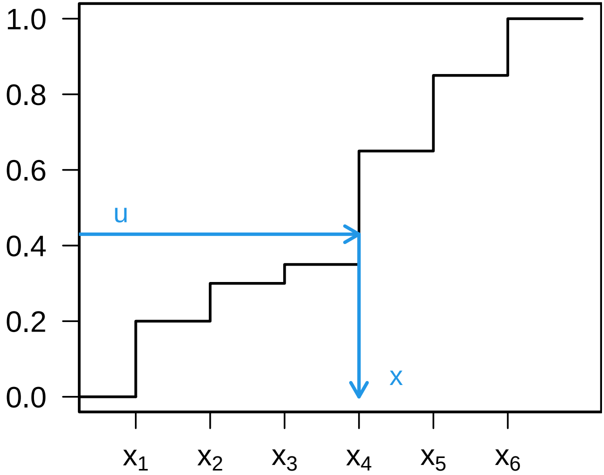
  if $u \in (F_{i-1}, F_i]$ **then**

    $x \leftarrow x_i$

  end if

end for

Review: sampling from discrete distribution (II)



Review: Probability integral transform to sample from continuous distributions

The inversion method (or probability integral transform approach) can be used to generate samples from an arbitrary continuous distribution with density $f(x)$ and CDF $F(x)$:

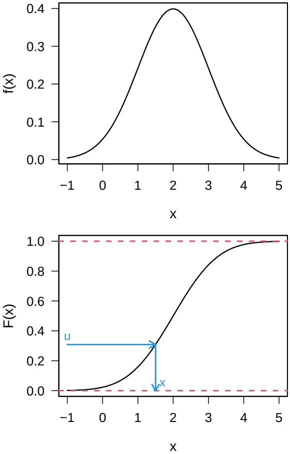
1. Generate random variable U from the standard uniform distribution in the interval $[0, 1]$.
2. Then is

$$X = F^{-1}(U)$$

   a random variable from the target distribution.

Probability integral transform to sample from continuous distributions

Let X have density $f(x)$, $x \in \mathbb{R}$ and CDF $F(x) = \int_{-\infty}^{x} f(z)dz$:



Simulation algorithm:

$u \sim U[0, 1]$

$x = F^{-1}(u)$

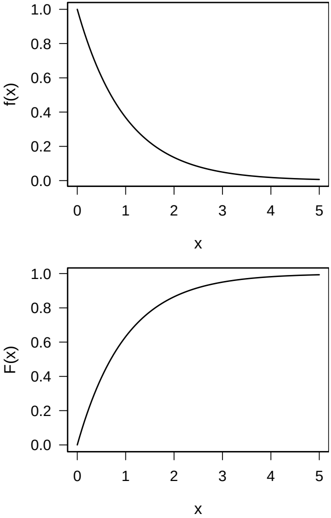
return x

Plan for today

Sampling from continuous distribution

- PIT transform
- Use relationship between random variable
  - Gamma distribution, χ^2 distribution
  - Linear transformation
  - Change of variables
- Bivariate techniques
  - Box-Muller algorithm (Normal distribution)
- Ratio of uniform method

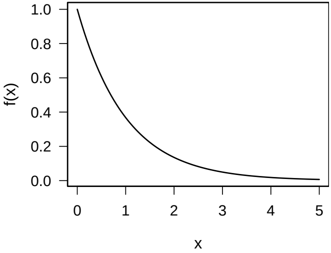
Example - Exponential Distribution

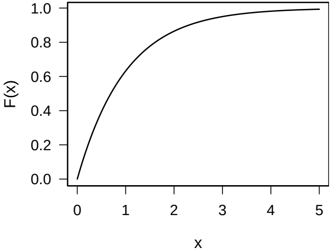


$$f(x) = \lambda \exp(-\lambda x) : \; x > 0$$

$$F(x) = 1 - \exp(-\lambda x)$$

Example - Exponential Distribution





$$f(x) = \lambda \exp(-\lambda x) : \; x > 0$$

$$F(x) = 1 - \exp(-\lambda x)$$

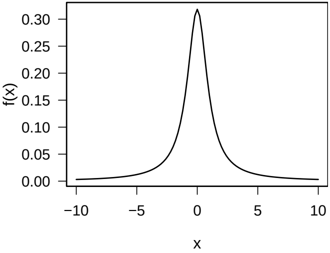
Simulation algorithm:

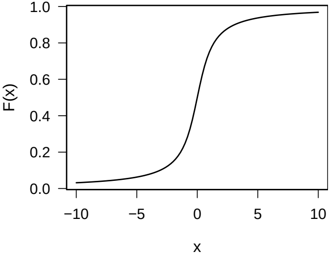
$u \sim U[0, 1]$

$x = -\frac{1}{\lambda} \log(u)$

return x

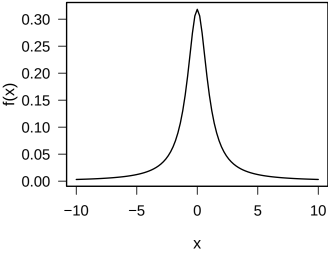
Example - Standard Cauchy distribution

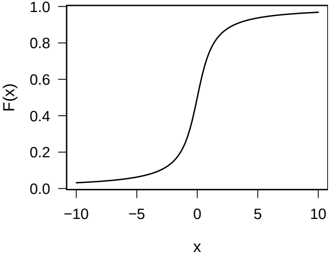




$$f(x) = \frac{1}{\pi} \cdot \frac{1}{1 + x^2}$$
$$F(x) = \frac{1}{2} + \frac{\arctan(x)}{\pi}$$
$$F^{-1}(y) = \tan\left[\pi\left(y - \frac{1}{2}\right)\right]$$

Example - Standard Cauchy distribution





$$f(x) = \frac{1}{\pi} \cdot \frac{1}{1+x^2}$$

$$F(x) = \frac{1}{2} + \frac{\arctan(x)}{\pi}$$

$$F^{-1}(y) = \tan\left[\pi\left(y - \frac{1}{2}\right)\right]$$

Simulation algorithm:

$u \sim U[0,1]$

$x = \tan[\pi(u - \frac{1}{2})]$

return x

Review: inverse transform technique

Let F be a distribution, and let $U \sim \mathcal{U}[0,1]$.

a) Let F be the distribution function of a random variable taking non-negative integer values. The random variable X given by

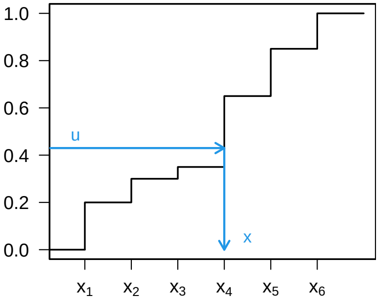
$$X = x_i \quad \text{if and only if} \quad F_{i-1} < u \leq F_i$$

has distribution function F.

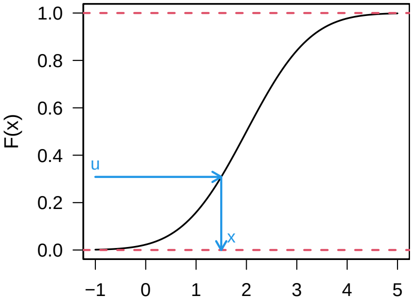
b) If F is a continuous function, the random variable $X = F^{-1}(u)$ has distribution function F.

Review: inverse transform technique (II)

a) Discrete case:



b) Continuous case:



The inverse transform technique is conceptually easy, but

- in the discrete case, a large number of comparisons may be necessary.
- in the continuous case, F^{-1} must be available.

Note

- The inversion method cannot always be used! We must have a formula for $F(x)$ and be able to find $F^{-1}(u)$. This is for example not possible for the normal, χ^2, gamma and t-distributions.
- In some cases we can use known relationships between RV to simulate

Using known relationships - Gamma distribution

Let $X \sim \text{Ga}(\text{shape}{=}\alpha, \text{rate}{=}\beta)$, i.e.

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta \cdot x}, x > 0.$$

If $X_1, \ldots, X_n \overset{\text{iid}}{\sim} \text{Exp}(\beta)$, then $Y = \sum_{i=1}^n X_i \sim \text{Ga}(n, \beta)$.

This gives how to simulate when α is an integer.

$y = 0$

for $i = 1, 2, \ldots, n$ **do**

  generate $u \sim U(0, 1)$

  $x \leftarrow -\frac{1}{\lambda} \log(u)$

  $y \leftarrow y + x$

end for

return y

Using known relationships - χ^2 distribution

Remember: $\chi^2_\nu = \text{Ga}(\frac{\nu}{2}, \frac{1}{2})$,

$X_1, \ldots, X_n \overset{\text{iid}}{\sim} \mathcal{N}(0,1) \Rightarrow \sum_{i=1}^n X_i^2 \sim \chi^2_n$.

Thus, we can simulate $X \sim \text{Ga}(\frac{n}{2}, \frac{1}{2})$ by

$x = 0$

for $i = 1, 2, \ldots, n$ **do**

generate $y \sim \mathcal{N}(0,1)$ $\quad\triangleright$ Still have to learn how

$x \leftarrow x + y^2$

end for

return x

Scale and location parameters

In Ga(α, β), β is a rate (inverse scale) parameter

$$X \sim \text{Ga}(\alpha, 1) \qquad \Leftrightarrow \qquad X/\beta \sim \text{Ga}(\alpha, \beta)$$

This gives us a way to sample from a Gamma distribution Ga($\frac{n}{2}, \beta$) where n is an integer

Gamma distribution - simulate $X \sim \text{Ga}(\frac{n}{2}, \beta)$

$x = 0$

for $i = 1, 2, \ldots, n$ **do**

  generate $y \sim \mathcal{N}(0, 1)$ ▷ Still have to learn how

  $x \leftarrow x + y^2$

end for

$x \leftarrow x$ ▷ $\text{Ga}(\frac{n}{2}, \frac{1}{2}), \chi^2_n$

$x \leftarrow \frac{1}{2}x$ ▷ $\text{Ga}(\frac{n}{2}, 1)$

$x \leftarrow \frac{1}{\beta}x$ ▷ $\text{Ga}(\frac{n}{2}, \beta)$

return x

Linear transformations

Many distributions have scale parameters, for example

$$X \sim \mathcal{N}(0,1) \quad \Leftrightarrow \quad \sigma X \sim \mathcal{N}(0,\sigma^2)$$

$$X \sim \mathsf{Exp}(1) \quad \Leftrightarrow \quad \frac{1}{\lambda} X \sim \mathsf{Exp}(\lambda)$$

$$X \sim \mathcal{U}[0,1] \quad \Leftrightarrow \quad \beta X \sim \mathcal{U}[0,\beta]$$

Adding a constant can also help in some situations

$$X \sim \mathcal{N}(0,1) \quad \Leftrightarrow \quad X + \mu \sim \mathcal{N}(\mu,1)$$

and thereby

$$X \sim \mathcal{N}(0,1) \quad \Leftrightarrow \quad \sigma X + \mu \sim \mathcal{N}(\mu,\sigma^2)$$

More general than just linear transformation: Change of variable

let $X \sim f_X(x)$ and $Y = g(X)$ with $g(\cdot)$ being a one-to-one function so that $Y = g^{-1}(X)$, then:

$$f_Y(y) = f_X(g^{-1}(x))|\frac{d\ g^{-1}(x)}{d\ x}|$$

Example: Change of variables

$X \sim \text{Exp}(1)$. We are interested in $Y = \frac{1}{\lambda}X$, i.e. $y = g(x) = \frac{1}{\lambda}x$.

$$g^{-1}(y) = \lambda y \qquad\qquad \frac{dg^{-1}(y)}{dy} = \lambda$$

Application of the change of variables formula leads to:

$$f_Y(y) = \exp(-\lambda y)\lambda$$

It follows: $Y \sim \text{Exp}(\lambda)$.

Exercise: Check other transformations, we mentioned.

Summary

- We can use know relationship between RV to derive samples from a RV we cannt sample directly from.
- If we can simulate from X and we know that $Y = g(X)$ and $g(\cdot)$ is invertible, then we can also get samples from Y
- Location and scale parameter are examples of linear transformation

Bivariate techniques

Remember:

$$\text{If } (x_1, x_2) \sim f_X(x_1, x_2)$$

$$\text{and } (y_1, y_2) = g(x_1, x_2)$$

$$\Updownarrow$$

$$(x_1, x_2) = g^{-1}(y_1, y_2)$$

where g is a one-to-one differentiable transformation. Then

$$f_Y(y_1, y_2) = f_X(g^{-1}(y_1, y_2))|\mathrm{J}|$$

with the determinant of the Jacobian matrix J

$$|\mathrm{J}| = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_2}{\partial y_1} \\ \frac{\partial x_1}{\partial y_2} & \frac{\partial x_2}{\partial y_2} \end{vmatrix}$$

$\Rightarrow$ Multivariate version of the change-of-variables transformation

Bivariate techniques (II)

If we know how to simulate from $f_X(x_1, x_2)$ we can also simulate from $f_Y(y_1, y_2)$ by

$$(x_1, x_2) \sim f_X(x_1, x_2)$$
$$(y_1, y_2) = g(x_1, x_2)$$

Return (y_1, y_2).

Example: Normal distribution (Box-Muller)

see blackboard

Review: Box-Muller algorithm

Generate

$x_1 \sim U(0, 2\pi)$

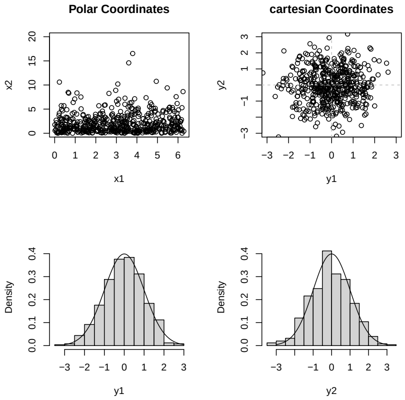
$x_2 \sim \exp(0.5)$

Compute

$y_1 \leftarrow \sqrt(x_2)\cos(x_1)$

$y_2 \leftarrow \sqrt(x_2)\sin(x_1)$

return (y_1, y_2)



Ratio-of-uniforms method

All the techniques seen untill now to sample from $f(x)$ require that we know the normalising constant of $f(x)$.

In many cases this is not the case. Often we only know that:

$$f(x) = \frac{1}{c} f^*(x)$$

where $f^*(x)$ is known while the constant (wrt x) c is unknown and is such that:

$$\int_{\mathcal{R}} f(x)dx = \frac{1}{c} \int_{\mathcal{R}} f^*(x)dx = 1$$

The Ratio of uniform method is a general method for arbitrary densities f known up to a proportionality constant.

Ratio-of-uniforms method

Theorem

Let $f^\star(x)$ be a non-negative function with $\int_{-\infty}^{\infty} f^\star(x)dx < \infty$. Let $C_f = \{(x_1, x_2) \mid 0 \leq x_1 \leq \sqrt{f^\star\left(\frac{x_2}{x_1}\right)}\}$.

a) *Then C_f has a finite area*

Let (x_1, x_2) be uniformly distributed on C_f.

b) *Then $y = \frac{x_2}{x_1}$ has a distribution with density*

$$f(y) = \frac{f^\star(y)}{\int_{-\infty}^{\infty} f^\star(u)du}$$

Example: Standard Cauchy distribution

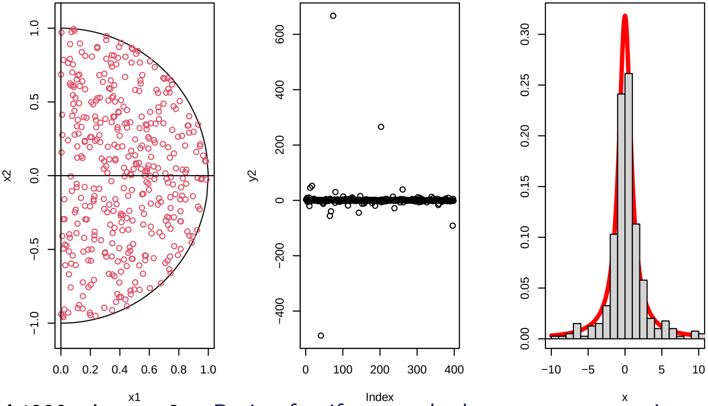
see blackboard

Algorithm to sample form a standard Cauchy

Generate (x_1, x_2) from $\mathcal{U}(C_f)$ ▷ How can we do this?

Compute $y = \frac{x_2}{x_1}$

return y



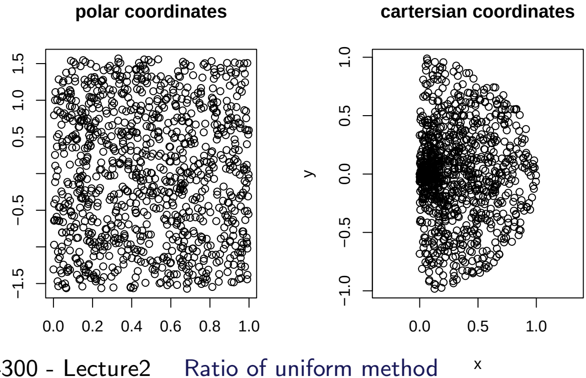
Sampling from the unit half circle

Idea: can we use polar coordinates?

$$x = u * cos(\theta)$$

$$y = u * sin(\theta)$$

can we use $\theta \sim \mathcal{U}(-\pi/2, \pi/2)$ and $u \sim \mathcal{U}(0, 1)$?



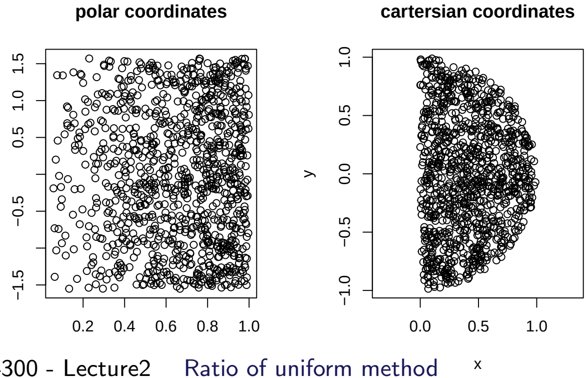
Sampling from the unit half circle

Idea: can we use polar coordinates?

$$x = u * cos(\theta)$$

$$y = u * sin(\theta)$$

Need to have $\theta \sim \mathcal{U}(-\pi/2, \pi/2)$ and $u^2 \sim \mathcal{U}(0, 1)$?



Proof of theorem

see blackboard

Ratio of uniform method

In general it can be hard to sample uniformly from C_f!!

It can be simplified under some contitions:

Theorem

Let $f^\star(x)$ be a non-negative function with $\int_{-\infty}^{\infty} f^\star(x)dx < \infty$. Let $C_f = \{(x_1, x_2) \mid 0 \leq x_1 \leq \sqrt{f^\star\left(\frac{x_2}{x_1}\right)}\}$.

If $f^\star(x)$ and $x^2\ f^\star(x)$ are bounded then $C_f \in [0, a] \times [b_-, b_+]$ with:

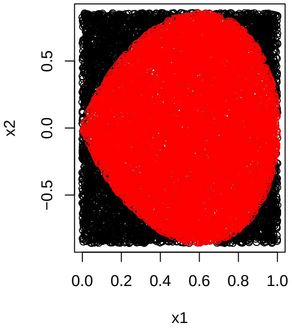
- $a = \sqrt{\sup_x f^\star(x)}$
- $b_- = -\sqrt{\sup_{x\leq 0} x^2\ f^\star(x)}$
- $b_+ = +\sqrt{\sup_{x>0} x^2\ f^\star(x)}$

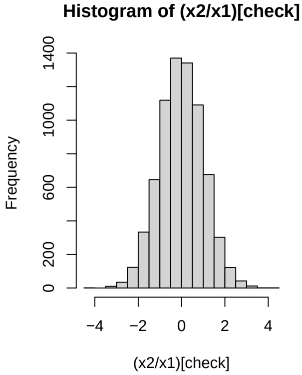
Proof of theorem

see blackboard

Ratio of uniform method: Simplification

- Rather than sampling uniformly from C_f, we can instead sample $(x_2\ x_2)$ uniformly from a rectangle containing C_f
- Reject sample if $(x_1\ x_2) \notin C_f$





Example: Normal distribution

see blackboard

Temporary page!

LaTeX was unable to guess the total number of pages correctly. A
there was some unprocessed data that should have been added t
the final page this extra page has been added to receive it.
If you rerun the document (without altering it) this surplus page
will go away, because LaTeX now knows how many pages to expe
for this document.