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

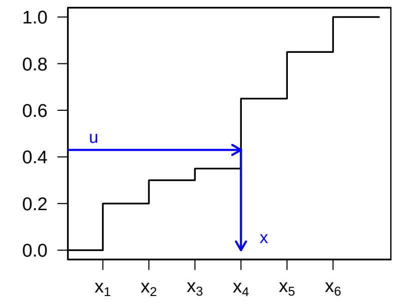
$$X = x_i \quad \text{if and only if} \quad F_{i-1} < U \leq F_i$$

has distribution function F.

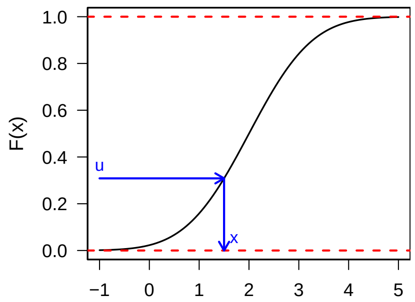
b) If F is a continuous function, the random variable $X = F^{-1}(U)$ has distribution function F.

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.

Bivariate techniques

Remember:

$$\begin{aligned} \text{If } (x_1, x_2) &\sim f_X(x_1, x_2) \\ \text{and } (y_1, y_2) &= g(x_1, x_2) \\ &\Updownarrow \\ (x_1, x_2) &= g^{-1}(y_1, y_2) \end{aligned}$$

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

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

with the determinant of the Jacobian matrix $\mathbf{J}$

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

$\Rightarrow$ 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

$$\begin{aligned} (x_1, x_2) &\sim f_X(x_1, x_2) \\ (y_1, y_2) &= g(x_1, x_2) \end{aligned}$$

Return (y_1, y_2).

Example: Normal distribution (Box-Muller)

see blackboard

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 $\mathcal{U}(C_f)$

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

return y.

Proof of theorem

see blackboard

Methods based on mixtures

Remember: $f(x_1, x_2) = f(x_1|x_2)f(x_2)$

Thus: To generate $(x_1, x_2) \sim f(x_1, x_2)$ we can

- generate $x_2 \sim f(x_2)$
- generate $x_1 \sim f(x_1|x_2)$

Note: This mechanism automatically provides a value x_1 from its marginal distribution, i.e. $x_1 \sim f(x_1) = \int_{-\infty}^{\infty} f(x_1, x_2) dx_2$.

$\Rightarrow$ We are able to generate a value for x_1 even when its marginal density is awkward to sample from directly.

Example: Simulation from Student-t (I)

The density of a Student t distribution with $n > 0$ degrees of freedom, mean μ and scale σ^2 is

$$f_t(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{1}{2}\right)} \frac{1}{\sqrt{n\sigma^2}} \left[1 + \frac{1}{n}\left(\frac{x-\mu}{\sigma}\right)^2\right]^{-\frac{n+1}{2}}, \quad -\infty < x < \infty.$$

Let

$$x_2 \sim \text{Ga}\left(\frac{n}{2}, \frac{nS}{2}\right)$$
$$x_1|x_2 \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{x_2}\right)$$

It can be shown that then

$$x_1 \sim t_n(\mu, S\sigma^2) \quad \text{(show yourself)}$$

Example: Simulation from Student-t (II)

Thus, we can simulate $x_1 \sim t_n(\mu, \sigma^2)$ by

$$x_2 \sim \text{Ga}\left(\frac{n}{2}, \frac{n}{2}\right)$$
$$x_1 \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{x_2}\right)$$

return x_1.

Another application is sampling from a mixture distribution, i.e. mixture of two normals.

Multivariate normal distribution

$\boldsymbol{x} = (x_1, \ldots, x_d)^\top \sim \mathcal{N}_d(\boldsymbol{\mu}, \Sigma)$ if the density is

$$f(\boldsymbol{x}) = \frac{1}{(2\pi)^{\frac{d}{2}}} \cdot \frac{1}{\sqrt{|\Sigma|}} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^\top \Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right)$$

with

- $\boldsymbol{x} \in \mathbb{R}^d$
- $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_d)^\top$
- $\Sigma \in \mathbb{R}^{d\times d}$, Σ must be positive definite.

Important properties (I)

Important properties of $\mathcal{N}_d(\boldsymbol{\mu}, \Sigma)$ (known from "Linear statistical models")

i) Linear transformations:
$\boldsymbol{x} \sim \mathcal{N}_d(\boldsymbol{\mu}, \Sigma) \Rightarrow \boldsymbol{y} = \mathbf{A}\boldsymbol{x} + \boldsymbol{b} \sim \mathcal{N}_r(\mathbf{A}\boldsymbol{\mu} + \boldsymbol{b}, \mathbf{A}\Sigma\mathbf{A}^\top)$, with $\mathbf{A} \in \mathbb{R}^{r\times d}$, $\boldsymbol{b} \in \mathbb{R}^r$.

ii) Marginal distributions:
Let $\boldsymbol{x} \sim \mathcal{N}_d(\boldsymbol{\mu}, \Sigma)$ with

$$\boldsymbol{x} = \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{bmatrix}, \quad \boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

Then

$$\boldsymbol{x}_1 \sim \mathcal{N}(\boldsymbol{\mu}_1, \Sigma_{11})$$
$$\boldsymbol{x}_2 \sim \mathcal{N}(\boldsymbol{\mu}_2, \Sigma_{22})$$

Important properties (II)

iii) Conditional distributions:
With the same notation as in ii) we also have

$$\boldsymbol{x}_1 | \boldsymbol{x}_2 \sim \mathcal{N}(\boldsymbol{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\boldsymbol{x}_2 - \boldsymbol{\mu}_2), \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})$$

iv) Quadratic forms:
$\boldsymbol{x} \sim \mathcal{N}_d(\boldsymbol{\mu}, \Sigma) \Rightarrow (\boldsymbol{x}-\boldsymbol{\mu})^\top \Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu}) \sim \chi^2_d$

Simulation from the multivariate normal

How can we simulate from $\mathcal{N}_d(\boldsymbol{\mu}, \Sigma)$?

Let $\boldsymbol{x} \sim \mathcal{N}_d(\mathbf{0}, \mathbf{I})$

$$\boldsymbol{y} = \boldsymbol{\mu} + \mathbf{A}\boldsymbol{x} \quad \overset{\text{i)}}{\Rightarrow} \quad \boldsymbol{y} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{A}\mathbf{A}^\top)$$

Thus, if we choose $\mathbf{A}$ so that $\mathbf{A}\mathbf{A}^\top = \Sigma$ we are done.

Note: There are several choices of $\mathbf{A}$. A popular choice is to let $\mathbf{A}$ be the Cholesky decomposition of Σ.