Inverse Transform Algorithms

Suppose $u \sim U(0,1)$. Then the rv X constructed as

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

has distribution F.

Example- Generate r.v. w/ distribution

$$F(x) = x^n \,, \quad 0 \le x \le 1$$

generate $u = F(x) = x^n$, or equivalently, $x = u^{1/n}$.

So,

(i) $u \sim U[0,1]$

(ii) compute $u^{1/n}$

(iii) This has cdf $F(x)$

Example- Generate exponentially distributed rv w/ rate λ:

$$F(x) = 1 - e^{-\lambda x}$$

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

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

So,

(i) $u \sim U[0,1]$

(ii) compute $-\frac{1}{\lambda}\log(1-u)$

(iii) This has cdf $F(x)$

$1-u$ has same distribution of u!

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

How to check if simulation is correct?

Suppose $x_1, \ldots, x_n$ is a sequence of random numbers simulated from cdf F.

Choose sequence of grid points $Z = \{z_1, \ldots, z_m\}$.

Evaluate empirical cdf $F_n(z) = \frac{1}{n} \sum_{k=1}^{n} I(x_k \leq z) \ \forall z \in Z$

For large n, $F_n(z)$ should be close to $F(z)$ for $z \in Z$

Why Inverse Transforms Work

Result: Suppose $u \sim U[0,1]$. Then for cdf F, the rv $X = F^{-1}(u)$ has distribution F.

Proof: Let $F_X(x)$ denote distribution of X generated by above algorithm. We want to show $F_X = F$. But

$$F_X(x) = Pr(X \leq x) = Pr(F^{-1}(u) \leq x)$$

F a cdf means

$$a \leq b \Leftrightarrow F(a) \leq F(b)$$

So

$$\begin{aligned} F_X(x) &= Pr(F(F^{-1}(u)) \leq F(x)) \\ &= Pr(u \leq F(x)) \\ &= F(x) \qquad (\text{since } u \sim U[0,1]) \end{aligned}$$

Note: have $X \sim F$, then $F(X) \sim \text{Uniform}$

Vector Random Variables

A multivariate random variable assigns to each outcome $\omega \in \Omega$ of a probabilistic experiment, a vector $X(\omega) \in \mathbb{R}^n$.

Then the CDF is

$$F_X(x) = \Pr(X_1(\omega) \leq x_1, \ldots, X_n(\omega) \leq x_n)$$

n-variate PDF is

$$f_X(x) = \frac{\partial^n}{\partial x_1 \ldots \partial x_n} F_X(x)$$

Note: $\int_{\mathbb{R}^n} p_X(x)\, dx = 1$

Example: n-variate Gaussian

$$X \sim N(\mu, \Sigma)$$

$$\frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right), \quad x \in \mathbb{R}^n$$

Marginalization

Given $f_{X,Y}(x,y)$

$$f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)\, dy$$

Given $F_{X,Y}(x,y)$

$$F_X(x) = F_{X,Y}(x, \infty)$$

Random Processes

A random process $X_k(\omega)$, $\omega \in \Omega$ is a family of random variables indexed by discrete time $k = 0, 1, \ldots, T$.

__Remarks__

1. Each outcome $\omega \in \Omega$ yields sample path $X_k(\omega)$
   (i) Fixing ω yields a deterministic function of time
   (ii) Fixing time n yields a random variable
   (iii) Fixing both n and ω yields a constant

2. Image Processing: spatial random processes

Example - IID Processes

X_k is independent and identically distributed if

$$p_{1,2,\ldots,n+1}(X_{n+1} = x \mid x_1, \ldots, x_n) = p_{n+1}(X_{n+1} = x)$$

"no memory"

and

$$p(X_{n+1} = x) \text{ has same pdf/pmf } \forall n$$

Most real-world random processes are non-iid and have memory

- speech, stock market, coordinates of vehicle
- linear difference equation
  - ↳ suppose V_k iid. Consider

$$X_{k+1} = \alpha X_k + V_k$$

then $\{X_k\}$ is NOT iid!

So, integral of iid process is NON-iid

Statistics of Random Processes

1. Probability Distribution: For random outcome ω at time k

$$F(x, k) = \Pr(X_k(\omega) \le x) \quad \forall x \in \mathbb{R}$$

2. PDF

$$f(x,k) = \frac{\partial}{\partial x} F_x(x,k)$$

PMF

$$f(x,k) = Pr(X_k(\omega) = x)$$

Clearly $f(x,k) \perp\!\!\!\perp k$ for iid processes

3. Second Order Distribution

$$F(x_1, x_2; k_1, k_2) = Pr(X_{k_1} \leq x_1, X_{k_2} \leq x_2)$$

4. Second Order Density

$$f(x_1, x_2; k_1, k_2) = \frac{\partial^2}{\partial x_1 \partial x_2} F(x_1, x_2; k_1, k_2)$$

Vector Random Process

1. Expected Value

$$\mu_k \triangleq \mathbb{E}[X_k] = \int_{-\infty}^{+\infty} x f_x(x,k)\, dx$$

$\mu_k \in \mathbb{R}^n$ is a deterministic time evolving vector

2. Variance, Auto-Covariance, Auto-Correlation

scalar $\quad \mathrm{Var}(X_k, X_l) = \mathbb{E}[(X_k - \mu_k)(X_l - \mu_l)]$

symmetric nxn matrix $\quad \mathrm{Cov}(X_k, X_l) = \mathbb{E}[(X_k - \mu_k)(X_l - \mu_l)^T]$

Note: $\mathrm{Var}(X_k, X_l) = \mathrm{trace}(\mathrm{Cov}(X_k, X_l))$

Statistical Inference

We relate probability models to the world in 2 ways:

(i) Law of Large Numbers

(ii) Central Limit Theorem

Law of Large Numbers

LLN relates statistics (real world) with probability (mathematical model)

For an iid process $\{X_k(\omega)\}$ two averages can be computed:

1. Expected value for fixed n (ensemble avg)

$$\mathbb{E}[X_k(\omega)] = \sum_{i=1}^{M} x_i f_X(x_i) = \mu$$

constant b.c. iid

2. Real life sample path time average (statistic) for fixed ω given N observations. (In real life you live one sample path)

Correction Pg 24

$$\hat{\mu}_N(\omega) = \frac{1}{N}\left(X_1(\omega) + \dots + X_N(\omega)\right)$$

Result: For iid process as $N \to \infty$, $\hat{\mu}_N$ converges to μ

In other words, the statistics $\hat{\mu}_N$ computed from data converges to mean of random variable, μ, Computed from probabilistic model!

More Precisely: Kolmogorov Strong LLN

$$\Pr\left(\{\omega : \lim_{N\to\infty} \hat{\mu}_N(\omega) \neq \mu\}\right) = 0$$

i.e. $$\hat{\mu}_N \xrightarrow{N} \mu \quad \text{w.p. } 1$$

Meaning: Suppose $X_n(\omega) \in \{0, 1\}$ iid w.p p_1, p_2.
Then as $N \to \infty$, LLN implies

number of ONES in sequence $(X_1, \dots, X_n) \xrightarrow{N} Np_2$

number of ZEROS in sequence $(X_1, \dots, X_n) \xrightarrow{N} Np_1$

More generally, if iid $X_n \in \{x_1, \dots, x_M\}$ then as $N \to \infty$

No. of x_m in N length sequence $(X_1, \dots, X_N) \to Np_m$

Note that if the sequence is NOT iid then LLN may NOT hold.
Consider

$$X_n = X_{n-1}, \quad X_0 = \begin{cases} 0 & \text{w.p. } 0.8 \\ 1 & \text{w.p. } 0.2 \end{cases}$$

the time average $\hat{\mu}_N = 0$ OR 1 but $\mu \triangleq \mathbb{E}[X_n] = 0.2$.
Thus $\mu \neq \hat{\mu}_N \Rightarrow$ LLN does NOT hold.

For LLN to hold you generally need to forget your initial condition geometrically fast

A random process for which LLN holds is called "ergodic". So any iid process is ergodic.