IE605: Engineering Statistics

Lecture 08

Manjesh K. Hanawal

Previous Lecture:

- F-distributions
- Convergence of RVs
- Consistency
- Order Statistics

This Lecture:

- Generating Random Samples

Two methods

- Direct method
- Indirect method

Direct method: Continuous case

- Generate random samples which has continuous CDF F
- $U \sim Unif(0,1)$. Set $X = F^{-1}(U)$. Then X has CDF F.
- $P(X \leq x) = P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x)$.

Example: Generate samples $\sim Exp(\lambda)$. $F(x) = 1 - e^{-\lambda x}$

- $F^{-1}(u) = -\frac{1}{\lambda}\log(1-u)$
- Set $X = -\frac{1}{\lambda}\log(1-U) \sim Exp(\lambda)$

Generating other continuous random variables

Gamma distribution *Gamma*(n, λ): For some interger n, $\lambda > 0$

- For $X_1, X_2, \ldots, X_n$ are iid and $\sim Exp(\lambda)$, we know $\sum_{i=1}^{n} X_i \sim Gamma(n, \lambda)$
- $U_1, U_2, \ldots, U_n$ are iid. with $\sim Unif(0, 1)$.
- $X_i = -\frac{1}{\lambda}\log(1 - U_i)$. $X_i \sim Exp(\lambda)$
- $X = \sum_{i=1}^{n} X_i = -\frac{1}{\lambda}\sum_{i=1}^{n}\log(1 - U_i)$.

Chi square distribution χ^2_{2n}: For some interger n

- $X \sim \chi^2_{2n} = Gamma(n, 1/2)$

Beta distribution *beta*(m, n): For some integers m, n

- $U_1, U_2, \ldots, U_m, U_{m+1}, \ldots, U_{m+n}$ are iid $\sim Unif(0, 1)$
- $X \sim \frac{\sum_{i=1}^{m}\log U_i}{\sum_{j=1}^{m+n}\log U_j} \sim beta(m, n)$.

Limitation: We cannot generate Chi square method with odd degree of freedom. Particularly, Gaussian distribution!

Other direct method

$$u = F(x) \implies u = \int_{\infty}^{x} f(z)dz$$

To invert we need to solve integration for each u!

Box-Muller method

Let U_1 and U_2 are iid with $\sim Unif(0,1)$. Define

$$R = \sqrt{-2\log U_1} \text{ and } \theta = 2\pi U_2.$$

Then

$$X_1 = R\cos(\theta) \text{ and } X_2 = R\sin(\theta)$$

are iid with $\sim \mathcal{N}(0,1)$. (Verify!)

Discrete case:

X takes discrete values $x_1, x_2, \cdots$. Assume $x_1 < x_2 < x_3, \cdots$.
To generate X satisfying CDF F.

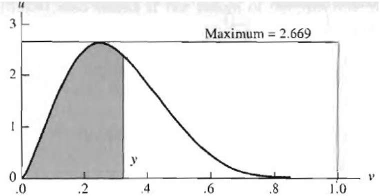
$$P(F(x_i) < U \leq F(x_{i+1})) = F(x_{i+1}) - F(x_i) = P(x_{i+1}).$$

Example:Generate $X \sim Bin(4, 5/8)$.
Generate $U \in Unif(0,1)$ and set

$$X = \begin{cases} 0 & if \quad 0 < U \leq 0.02 \\ 1 & if \quad 0.02 < U \leq .152 \\ 2 & if \quad .152 < U \leq .481 \\ 3 & if \quad .481 < U \leq 0.847 \\ 4 & if \quad .847 < U \leq 1 \end{cases}$$

Indirect method

Consider a distribution which takes value in $[0, 1]$.



Let (U, V) are iid $\sim Unif(0, 1)$. f is the given PDF. $c = \max_y f(y)$

$$P\left(V \leq y, U \leq \frac{1}{c}f(V)\right) = \int_0^y \int_0^{f_X(v)/c} dudv$$

$$= \frac{1}{c}\int_0^y f_X(v)dv = \frac{1}{c}P(X \leq y)$$

Set $y = 1$. We get $P\left(U \leq \frac{1}{c}f(V)\right) = \frac{1}{c}$

Indirect method contd..

$$P(X \le y) = \frac{P\left(V \le y, U \le \frac{1}{c} f(V)\right)}{P\left(U \le \frac{1}{c} f(V)\right)}$$
$$= P\left(V \le y, U \le \frac{1}{c} f(V) | U \le \frac{1}{c} f(V)\right)$$

Algorithm:

1. (a) Generate (U, V) iid $Unif(0, 1)$
2. if $U \le \frac{1}{c} f(V)$. Set $X = V$
3. else go to step 1.

Accept/Reject Algorithm

Remove the requirement that distribution takes value in $[0, 1]$.

Let f be any given distribution and V is a RV with pdf f_V

$$M = \max_y \frac{f(y)}{f_V(y)} < \infty (\text{assume!})$$

Algorithm

1. Generate $U \sim Unif(0, 1)$ and $V \sim f_V$ independently
2. If $U \leq \frac{1}{M} f(V)$, set $X = V$
3. Else, go to setp 1.

Calim: X has pdf f.

If M is not bounded . Accept/Reject algorithm will not work.
Metropolis Algorithm extends the ideas.